Mg^{2+} Mediated Sequence-Specific Binding of Transcriptional Activator Protein C of Bacteriophage Mu to DNA†

Amitabha De,‡ V. Ramesh,§ S. Mahadevan, and V. Nagaraja*

Microbiology and Cell Biology Department, Indian Institute of Science, Bangalore 560 012, India

Received September 2, 1997; Revised Manuscript Received December 22, 1997

ABSTRACT: The contributions from the secondary structure of the transcriptional activator protein C of bacteriophage Mu to its specific DNA binding and the influence of various factors, viz., electrolytes, and minor groove and major groove binders on this protein−DNA interaction have been addressed. Circular dichroism (CD) spectral results suggest that, in the absence of Mg^{2+}, C protein exhibits a β-pleated sheetlike structure and Mg^{2+} changes the conformation to a more α-helical structure which could provide specific geometrical constraints complementary to those of DNA helix. Thus, Mg^{2+} acts as a cofactor for the binding of the C protein to its specific site in DNA by inducing conformational changes in the protein. Competitive binding studies with minor and major groove binding drugs, viz., distamycin A and methyl green, respectively, and the DMS footprinting data indicate that the C protein recognizes the major groove of DNA during complex formation. Further, upon major groove binding, C protein brings about changes in DNA conformation; such conformational changes could have implications in the transcription process.

The mutual recognition of molecules is decisive in enzyme catalysis and inhibition, gene expression and its control, DNA replication, immune response, and drug action. The mechanisms of mutual recognition between DNA and proteins are central to many of the vital processes in structural molecular biology. DNA binding proteins are involved in a number of basic cellular processes, including transcription, DNA replication, transposition, restriction, recombination, and DNA repair. The necessary complementarity between the participant proteins and DNAs may either preexist or arise during the process of complex formation. In the latter cases, it is DNA that usually contributes to the specificity of protein interactions by its ability to exist in alternate conformations or to deform its structure to accommodate protein binding, e.g., interaction of DNA with RNA polymerase, *Eco*RI restriction endonuclease (*1*, *2*). However, examples where more dramatic changes occur in the protein have also been reported, e.g., leucine zipper protein−DNA complexes (*3*). Thus, the simultaneous interaction of a number of functional groups on the two partners must be involved, and these groups could be positioned, at least temporarily, in a specific conformational relationship which permits the favorable interaction between DNA and proteins.

We address here some of the factors influencing the sequence-specific binding of transcriptional activator protein C of bacteriophage Mu to its site in *mom* promoter region. The *mom* gene of bacteriophage Mu, which encodes a unique DNA modification function, is the last gene to be expressed in the lytic cycle of the phage and is subjected to a complex mode of regulation (*4*, *5*). The *mom* gene expression is modulated by two phage-encoded proteins C and Com (*6*). The C protein, a middle gene product of bacteriophage Mu, acts as a transcription activator at four late promoters, including P_{mom} (*7*, *8*). The C protein binding site has been mapped on *mom* promoter by MPE.Fe(II) footprinting analysis (*7*). The protected region is in the −35 element of the promoter. Two overlapping *Escherichia coli* RNA polymerase (RNAP)[1] binding sites, named P_1 and P_2, are present in the *mom* promoter region (*9*). In the absence of C protein, RNAP binds preferentially to the P_2 site, whereas in its presence it binds to the P_1 site as the C protein binding site overlaps with the P_2. To analyze structure−function relationships and specific interaction between the C protein and its cognate site, we have engineered the hyperexpression (*10*) of C protein using the T7 expression system. The purified recombinant protein is composed of 138 amino acids and binds to its cognate site with a very high equilibrium constant, on the order of 2.0×10^{12} M^{-1} (*10*, *11*). Gel filtration chromatography, heterodimer DNA binding assay, protein cross-linking, and deletion analysis show that the protein binds to DNA as a dimer and implicate a role for a helix−turn−helix (HTH) motif located toward the carboxyl-terminal region of the protein in DNA binding (*12*, *13*). To obtain an insight into the DNA binding events of the C protein, biophysical and biochemical techniques have been employed, and presented here is a discussion of the contribution of the structure of the protein to specific DNA binding.

† This work was supported by a grant from the Department of Science and Technology, Government of India. V.R. was supported by the Jawaharlal Nehru Centre for Advanced Scientific Research. A.D. was supported by the Department of Biotechnology Postdoctoral Training Program.

* Author to whom correspondence should be addressed at Microbiology and Cell Biology Department, Indian Institute of Science, Bangalore 560 012, India. Fax: 91-80-3341683. E-mail: vraj@cge.iisc.ernet.in.

‡ Present address: National Dairy Development Board, Anand 388 001, India.

§ Present address: Department of Biology, Massachussets Institute of Technology, Cambridge, MA 02139.

[1] Abbreviations: bp, base pair(s); CD, circular dichroism; DMS, dimethyl sulfate; EMSA, electrophoretic mobility shift assay; HTH, helix−turn−helix; PAGE, polyacrylamide gel electrophoresis; RNAP, RNA polymerase; TEMED, *N*,*N*,*N*′,*N*′-tetramethylethylenediamine.

S0006-2960(97)02171-5 CCC: $15.00
Published on Web 02/24/1998

MATERIALS AND METHODS

The restriction enzymes, Klenow fragment of DNA polymerase I, calf intestine alkaline phosphatase, and phage T_4 polynucleotide kinase were from Boehringer-Mannheim. [α-^{32}P]dATP (3000 Ci/mmol) and [γ-^{32}P]dATP (3000 Ci/mmol) were from Bhabha Atomic Research Centre; acrylamide, bisacrylamide, TEMED, ammonium persulfate, and other chemicals were obtained from Sigma.

Protein Purification. C protein was purified from *E. coli* strain BL21(DE3) harboring plasmid pVR6 as described (*11*). The protein eluted from the immunoaffinity column was dialyzed against buffer A [20 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 3% (v/v) glycerol, 7 mM 2-mercaptoethanol]. The renatured protein was stored at −20 °C in buffer B [20 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 10% (v/v) glycerol, 7 mM 2-mercaptoethanol].

Labeling DNA Fragments. A 260 bp DNA fragment, a source for C protein binding site, was prepared and purified as described earlier (*10*). The plasmid pUW4 (100 μg) was digested with *Eco*RI−*Hin*dIII, and the fragments were separated by 1% agarose gel electrophoresis. The 260 bp fragment was extracted from the gel piece and purified by phenol and chloroform extractions followed by ethanol precipitation (*14a*). Recovered DNA was estimated by A_{260}, and purity was also tested by gel electrophoresis. Both of the 3′ ends were filled with [α-^{32}P]dATP and other dNTPs by using a Klenow fragment of DNA polymerase I. Specific activity of the end-labeled fragment was 75 cpm/pmol as measured by Cerenkov radiation.

Electrophoretic Mobility Shift Assay (EMSA). Binding of C protein to the DNA fragment was assayed by the electrophoretic mobility shift method. The assay was performed in EMSA buffer [20 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 90 mM KCl, 0.1 mM EDTA, 3% glycerol] unless mentioned otherwise. Reaction mixture (20 μL) containing 2.5 ng of labeled DNA in the absence or presence of varying amounts of distamycin A or methyl green and 50 ng of protein was incubated for 20 min at 25 °C. After incubation, 4 μL of loading buffer [20 mM Tris-HCl (pH 7.5), 10% glycerol] containing 0.01% bromophenol blue and 0.01% xylene cyanol was added. All the samples were analyzed on a 4% polyacrylamide gel (30:0.8) equilibrated and electrophoresed in 0.5× TBE [45 mM Tris-borate, 1 mM EDTA (pH 8.3)] buffer at 100 V for 5 h at 4 °C, dried, and subjected to autoradiography. The radioactivity in the bands was quantified by Cerenkov counting.

DNase I Footprinting. The purified C protein (100−500 ng) was incubated in a 20 μL reaction volume with an end-labeled 260 bp DNA fragment (1 pmol, 10^5 cpm) in the footprinting buffer [10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, 10 mM $MgCl_2$, 4 mM $CaCl_2$, 3% glycerol, 60 mM KCl, 2% PEG-8000, 1 mM DTT] for 10 min at 25 °C. DNase I (Worthington Biochemical Corp., 1828 U/mg) was freshly diluted in the same buffer (on ice) from a stock solution of 10 mg/mL [in storage buffer 20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 50% glycerol, 50 μg/mL BSA] to 1 ng/μL, and 1 μL was added to the reaction mixture. The reaction was performed for 40 s at the same temperature and terminated by the addition of 20 μL of stop buffer [50 mM Tris-HCl (pH 8.0), 15 mM EDTA, 0.3% SDS, 150 mM NaCl]. The volume was increased to 100 μL with TE [10 mM Tris-HCl (pH 7.4), 1 mM EDTA (pH 8.0)] and deproteinized by extraction with phenol and then chloroform. The DNA was precipitated with 3 volumes of ethanol and ammonium acetate in the presence of 5 μg of carrier tRNA.

Methylation Protection and Interference. A 25 bp double-stranded oligonucleotide corresponding to the C binding site in P_{mom} was used. For this purpose, the top-strand 25 base single-stranded oligonucleotide (5′-AGATCGATTATGCCCCAATAACCAC-3′) was synthesized and purified using standard procedures (*14b*). This top strand was mixed with the complementary strand (bottom) oligonucleotides and incubated at 90 °C for 2 min and then allowed to cool slowly to room temperature. The samples were run on native PAGE to assess the formation of duplex oligo along with the respective single-stranded oligonucleotides as markers. The individual oligonucleotides were purified, phosphorylated using [γ-^{32}P]ATP and T4 polynucleotide kinase, annealed, and used in methylation protection and interference experiments. For protection from methylation by dimethyl sulfate (DMS), the C protein−DNA complex (500 ng protein; 1 pmol of labeled DNA oligo, 10^5 cpm) was preformed in 100 μL volume in binding buffer [10 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 50 mM NaCl, 1 mM dithiothreitol, 5% glycerol] and equilibrated at 25 °C for 10 min. DMS was freshly diluted in binding buffer to a final concentration of 2%, and 10 μL was added to the preformed complex. After incubation for about 10 min, the reaction was stopped by the addition of 25 μL of DMS stop solution (1.5 M sodium acetate and 1 M 2-mercaptoethanol) and deproteinized by phenol and chloroform extractions, and the DNA was precipitated with 3 volumes of ethanol and 5 μg of carrier tRNA. The DNA was reprecipitated, washed with 70% ethanol, and suspended in 90 μL of sterile distilled water. The control samples were methylated under identical conditions in the absence of C protein.

Methylation interference was performed by methylating the DNA under identical conditions prior to complex formation. After removal of DMS, the DNA−protein complex was formed as described above, and the free and bound DNA were resolved on a 4% polyacrylamide gel. The bands were located by autoradiography, and the DNA was excised and eluted (*15*) and suspended in 90 μL of sterile distilled water.

The G sequencing ladders (*15*) were generated by methylating the labeled oligonucleotide under the following conditions. Between 1 and 10 pmol of labeled DNA in a 200 μL reaction volume [50 mM sodium cacodylate (pH 8.0), 1 mM EDTA] was incubated with 1 μL of DMS for 2 min at 25 °C. The reaction was terminated by the addition of 50 μL of DMS stop solution (mentioned above) and precipitated with 750 μL of ethanol and 5 μg of carrier tRNA. The DNA was reprecipitated, washed with 70% ethanol, and suspended in 90 μL of sterile distilled water. The strand cleavage reaction at the methylated guanines was performed by the addition of 10 μL of piperidine to 90 μL of methylated DNA and heating at 90 °C for 30 min followed by freeze-drying.

Fluorescence Measurements. Fluorescence measurements were carried out using a JASCO FP-777 spectrofluorimeter. In each case 10 mol of C protein solution (2 mL) was taken in buffer C [20 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$,

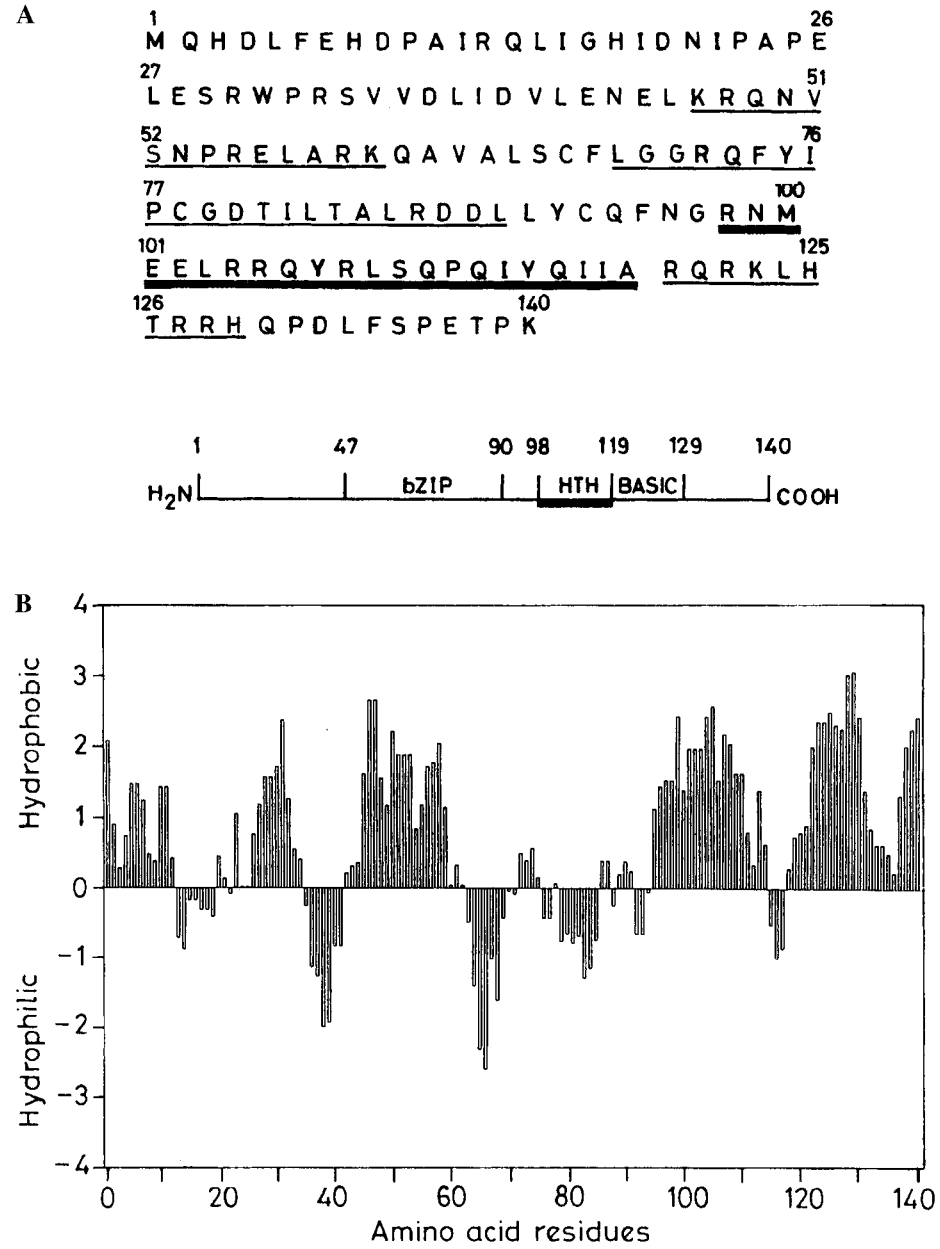

FIGURE 1: (A) Amino acid sequence of the C protein showing putative motifs. A single letter code is used. (B) Kyte−Doolittle hydropathy profile of C Protein.

90 mM KCl, 0.1 mM EDTA, 3% glycerol] in a quartz cuvette of 1 cm path length at 25 °C. In the case when 6 M urea was treated, preincubation was done for 30 min. Similarly, after addition of DNA to protein solution in buffer C, the reaction mixture was also set aside for 20 min for association. The dissociation of the DNA−protein complex in the presence of 500 mM NaCl was also performed in a similar way.

Freshly prepared 4 M stock solutions of KI, CsCl, and acrylamide were used for quenching. A C protein sample (100 μg/mL) was present in 2 mL of buffer C. Quenching studies were performed using excitation at 280 nm (5 nm band-pass) and emission at 330 nm (5 nm band-pass) for all quenchers. Fluorescence intensity values were corrected for each dilution. DNA binding studies were performed using the excitation at 300 nm to avoid inner filter effects, if any, that would arise from DNA which has significant absorption at 280 nm. The quenching constants for various quenchers were determined using the Stern−Volmer equation:

$$F_o/F = 1 + K_{sv}[Q]$$

where F_o and F are fluorescence intensities in the absence or presence of quencher. K_{sv} is the Stern−Volmer quenching constant, and [Q] is the concentration of quencher in the cell. The slope of the plot of F_o/F vs [Q] is taken as the measure of the quenching constant.

Circular Dichroism (CD) Measurements. CD spectra were recorded on a JASCO 500A spectropolarimeter using a path length of 0.1 cm at 25 °C. Samples were prepared at room temperature in buffer [20 mM Tris-HCl (pH 8.0), 100 mM KCl]. In cases where the effect of Zn^{2+} or Mg^{2+} were studied, the respective divalent ions are added to this buffer to a final concentration of 5 mM. The 220 bp DNA fragment containing the C binding site was prepared and purified as described for the 260 bp fragment, with the exception that

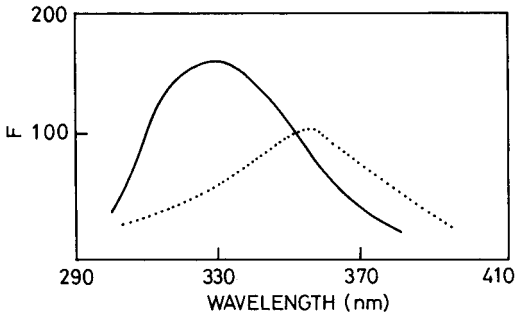

FIGURE 2: Emission spectrum of native protein in the absence (—) and in the presence (•••) of 6 M urea. Excitation at 280 nm.

the plasmid pUW4 was digested with *Eco*RI−*Bam*HI. An amount of 10 μM 220 bp DNA fragment bearing the C binding site was incubated with C protein (200 μg/mL) before CD measurements. Sonicated calf thymus DNA served as a nonspecific control. A mean amino acid residue weight of 98.9 was used for the calculation of the molar ellipticity values of the proteins (*16*). Interaction of groove binding ligands or interaction of C protein with DNA was also carried out at 25 °C in a cell of 0.1 cm path length. In the case of DNA−ligand interaction, at each concentration of ligand, the reaction mixture was set aside for a period of 30 min for equilibration, as was also done for a mixture of DNA and protein. A mixture of DNA, ligand, and protein was set aside for 30 min to attain equilibrium. The concentration of DNA used was calculated from A_{260}, and the concentrations of proteins were estimated according to Bradford (*17*).

RESULTS

The sequence of the transcriptional activator protein C of bacteriophage Mu is shown in Figure 1A. The 16.8 kDa protein has certain putative structural motifs. A single tryptophan (Trp) at position 31 is present in the motif Ser-Arg-Trp-Pro-Arg in the primary sequence of the protein. However, the Kyte−Doolittle hydropathy plot analysis showed that the tryptophan residue is buried in a hydrophobic pocket (Figure 1B). The location of Trp within the hydrophobic pocket was verified by an emission spectrum. Upon excitation at 280 nm, the resulting emission spectrum gives a peak around 330 nm. In proteins that contain tryptophan, both shifts in wavelength and changes in intensity are generally expected upon unfolding. The tryptophan emission of a native protein can be greater or smaller than the emission of free tryptophan in aqueous solution. Consequently, both increase and decrease in fluorescence intensity can occur upon protein unfolding. The emission maximum is usually shifted from shorter wavelengths to about 350 nm, which corresponds to the fluorescence maximum of tryptophan in aqueous solution. In a hydrophobic environment, such as in the interior of folded protein, tryptophan emission is expected to occur at shorter wavelengths. Emission spectra of the native C protein and of unfolded protein are shown in Figure 2. The emission maximum at 332 nm of the native protein was shifted to 346 nm on unfolding. Red-shift in the presence of urea indicates that the tryptophan moiety was buried in a hydrophobic pocket. A similar observation has been interpreted to confirm the inaccessibility of the single tryptophan residue to solvent in RNase T1 (*18*). To understand the nature of amino acids neighboring Trp, quenching studies were performed with neutral, cationic, and anionic quenchers. Figure 3 shows that the neutral quencher acrylamide quenches linearly up to 230 mM, whereas CsCl and KI quench linearly only up to a concentration of 60 mM. Beyond that concentration, downward deviation was observed in the latter two cases. The Stern−Volmer quenching constants have been calculated from the slopes, and their values have the following order: acrylamide > KI > CsCl. Further, the quenching constant in the presence of CsCl is half that of KI, indicating that the tryptophan was likely to be present in the local cationic atmosphere, within the hydrophobic pocket.

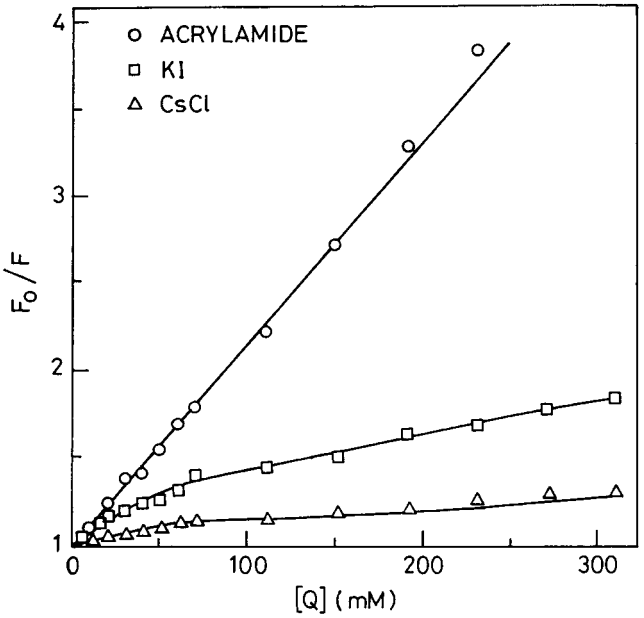

FIGURE 3: Quenching of tryptophan emission with acrylamide, KI, and CsCl. Stern−Volmer plots of the ratio of tryptophan fluorescence intensity in the absence (F_o) and presence (F) of acrylamide (○), potassium iodide (□), or cesium chloride (△) in buffer C at 25 °C.

Fluorescence experiments have also been carried out to study the interaction of the C protein with DNA, and the results of fluorescence measurements of protein−DNA complexes are shown in Figure 4A. In the presence of a specific DNA fragment containing the C binding site, 20% reduction in fluorescence is observed; this suggests the association of C protein with DNA. These experiments were carried out in the presence of 100 mM NaCl. During the course of the investigation, electrophoretic mobility shift assay at different NaCl concentrations shows that the protein−DNA complex is stable up to a concentration of the salt of 350 mM (Figure 5, see below), beyond which no complex formation was detected. This is confirmed by emission measurements at high salt conditions. Reaction mixtures containing C protein in the absence and presence of specific DNAs were separately treated with 500 mM NaCl. In both cases, the final emission spectra (Figure 4B) are similar, indicating thereby the dissociation of DNA−C protein complex at the higher concentration of NaCl. Further, it confirms that the quenching observed at lower NaCl concentration (100 mM) in the presence of DNA is due to its association with C protein. The above experiments were carried out in the presence of 5 mM Mg^{2+} and were repeated in Mg^{2+}-free buffer.

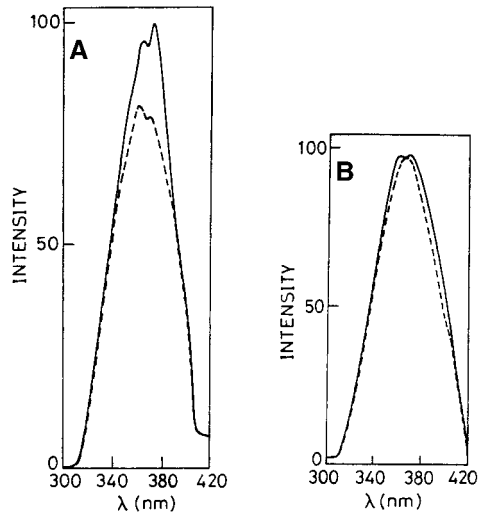

FIGURE 4: Association and dissociation of DNA−C protein complex. (A) Emission spectra of purified C protein (10 μM) in buffer C in the absence (—) and presence (---) of 10 μM DNA. (B) Emission spectra of the protein after supplementation of 500 mM NaCl in the absence (—) and presence (---) of 10 μM DNA. Excitation at 300 nm.

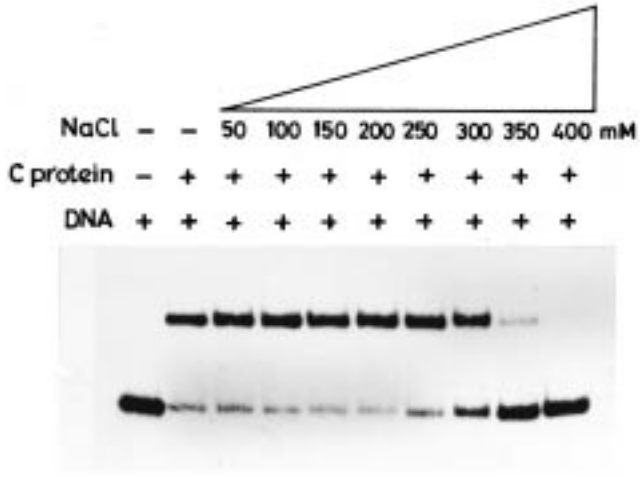

FIGURE 5: Effect of NaCl on C protein−DNA complex.

Quenching of the tryptophan emission by DNA was not observed in the absence of Mg^{2+} (data not shown), suggesting that the latter seems to be essential for the binding of C protein to DNA. The essential requirement and the role of Mg^{2+} in sequence-specific DNA binding of C protein are further addressed below by circular dichroism studies and electrophoretic mobility shift assays.

The influence of Mg^{2+} and other cations on DNA binding of C protein was studied further using EMSA. Figure 5 shows the effect of varying the concentration of NaCl on the protein−DNA complex. Beyond 350 mM NaCl concentrations, the complex dissociation is consistent with the emission results. The influence of divalent cations on complex formation was also tested. Only in the presence of 5 mM Mg^{2+} was effective complex formation observed (inset of Figure 6, see below). Weak binding was observed in the presence of Mn^{2+}, while in the case of Cd^{2+} or Zn^{2+}, no binding was detected. Thus, magnesium ion is essential for the binding of C protein to DNA.

There have been predictions concerning the type of protein secondary structure or motif that might bind to DNA. Protein

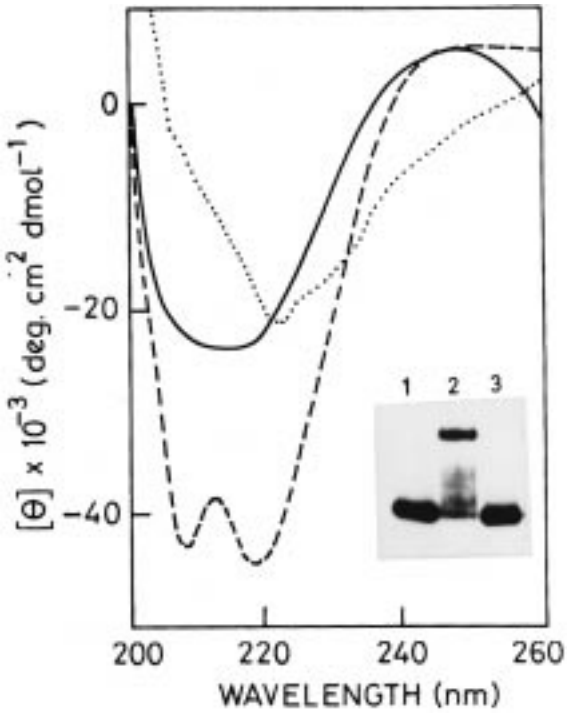

FIGURE 6: Requirement of divalent cation on C protein structure. Circular dichroic spectra of purified C protein (200 μg/mL) in 20 mM Tris-HCl (pH 8.0) buffer containing 100 mM KCl (—), 100 mM KCl + 5 mM Zn^{2+} (···), and 100 mM KCl + 5 mM Mg^{2+} (---). Inset: The effect of divalent cations on C protein−DNA complex. The divalent cations were supplemented to a final concentration of 5 mM. Lane 1, free DNA; lane 2, DNA + C protein with 5 mM Mg^{2+}; lane 3, DNA + C protein without Mg^{2+}.

α-helixes and β-pleated-sheet structures have rather specific geometrical properties and dimensions, some of which are complementary to those of the DNA helix. Immunoaffinity purified C protein is structurally intact and, therefore, a good substrate for structural analysis. As a preliminary measure of global structural content, the purified C protein was subjected to analysis by CD spectroscopy, which is a convenient and widely used method for studying the conformations and conformational changes of proteins under different conditions. The CD spectrum of C protein (Figure 6) in 20 mM Tris-HCl (pH 8.0) displayed a broad peak around 220−210 nm, indicating that the protein exists mainly as β-pleated sheet-like structure (*19*). The structural content of the protein was not changed significantly either in the presence of 100 mM KCl or by the subsequent addition of 5 mM Zn^{2+} to the salt-containing buffer. However, the effect of Mg^{2+} (5 mM) addition is more dramatic on the spectral pattern. There is a significant increase in the magnitude of the helix-associated CD bands at 220 and 208 nm, indicating that Mg^{2+} is essential for its ordered structure. The α-helical content was found to be 33%, which is based on the theoretical calculation of Chou and Fasman (*20*) using the PEPPLOT computer program. These conformational changes leading to increased α-helicity seem to be a prerequisite for DNA binding, as gel retardation experiments show the requirement of Mg^{2+} for binding of protein to DNA (inset of Figure 6).

Apart from the specific sequence, the wide and narrow grooves of the DNA target site provide opportunity for the protein to bind, and thus could be probed by specific nonintercalative ligands. We have used distamycin A and

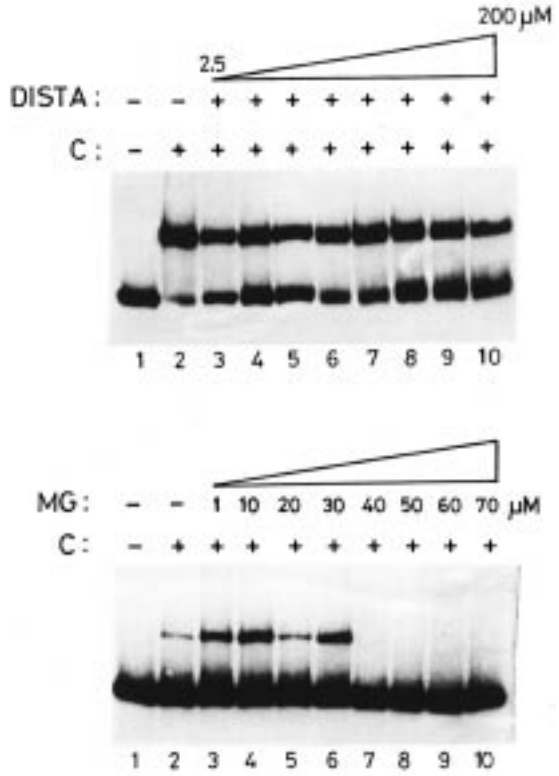

FIGURE 7: Effect of groove-specific ligands on DNA binding of C protein. Fifty nanograms of C protein and 2.5 ng of labeled DNA (10 000 cpm) were used in the assay. (A) Lane 1, no protein; lane 2, 50 ng of C protein; lanes 3−10 have methyl green at concentrations ranging from 1 to 70 μM, preincubated with DNA followed by addition of the protein. (B) Lane 1, no protein; lane 2, 50 ng of C protein; lanes 3−10 have distamycin A at concentrations ranging from 2.5 to 200 μM. The drug was preincubated with DNA before the addition of the protein.

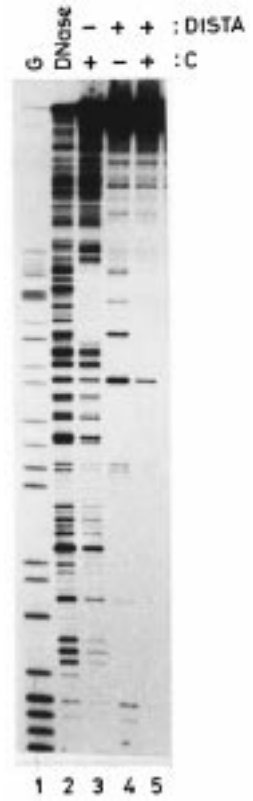

FIGURE 8: DNase I footprinting. Lanes: 1, G sequencing ladder; lane 2, free DNA; lane 3, C protein-bound DNA; lane 4, distamycin-bound DNA; lane 5, C protein-bound DNA in the presence of distamycin.

methyl green as groove-specific ligands to identify the role of structural features, narrow and wide grooves in particular, required for the specific protein−DNA interaction. Distamycin A, the pyrrole amide antibiotic, is a well-known DNA minor groove binder while methyl green binds to the wide major groove (*21*, *22*). The effect of the drugs on the DNA−C protein association was studied using electrophoretic mobility shift assay (EMSA), DNase I footprinting, and circular dichroism spectra. Under the experimental conditions, up to 30 μM concentrations, methyl green has no effect on the DNA−protein complex. However, beyond 40 μM concentrations, no protein−DNA complex was detected. (Figure 7A, lanes 7−10). Further, competition assay was performed keeping the concentrations of DNA and methyl green constant and increasing the concentration of protein. On gradually increasing the amount of C protein, the amount of complex formation increases (data not shown), indicating that methyl green and C protein are essentially competing for the binding site at the major groove. In case of distamycin A, the concentration was varied from 0 to 200 M. In this range distamycin A has no inhibitory effect on DNA−C protein association (Figure 7B). Further, no significant effect on protein−DNA complex formation was observed on preincubation of DNA with distamycin and subsequent addition of the protein. More information on the effect of distamycin A on C protein−DNA complex was obtained from DNase I footprinting. DNase I footprinting results (Figure 8) showed that C protein protects the 28−30 bp region in the 220 bp *mom* promoter DNA fragment. This fragment contains approximately 12 distamycin binding sites and among them two are present in the C binding site. Footprinting studies showed that the C protein binding site is not protected by distamycin in the presence of the protein, while distamycin protection of other sites is unaffected.

A sensitive tool for analyzing DNA conformation is the CD spectrum, which has been widely used as a probe to follow drug- and protein-induced structural changes in DNA. A number of studies involving drug−DNA interactions have helped to infer the role(s) of structural elements of DNA such as contacts in the wide and narrow grooves and sequence-dependent conformations in protein−nucleic acid interactions. The effect of distamycin on the CD spectra of DNA alone and of the DNA−protein complex has been studied. The drug itself has no ellipticity in the 200−360 nm region in the CD spectra. However, DNA-induced positive ellipticity was observed in the 310−330 nm range. For a particular concentration of DNA, positive ellipticity in this range increases with increase in concentration of distamycin (data not shown), attributable to the strong interaction of the latter to DNA, and the peak at 260 nm region remains unaltered in the presence of 0−200 μM distamycin A. The conformational changes in DNA on binding to the drug are not significant, as shown from the insignificant changes in the CD of the former. In the presence of a saturating amount of protein in the reaction mixture, DNA exhibits a decrease in positive ellipticity values in its CD, indicating that DNA undergoes a conformational change upon complex formation with C protein. Sonicated calf thymus DNA does not exhibit any significant

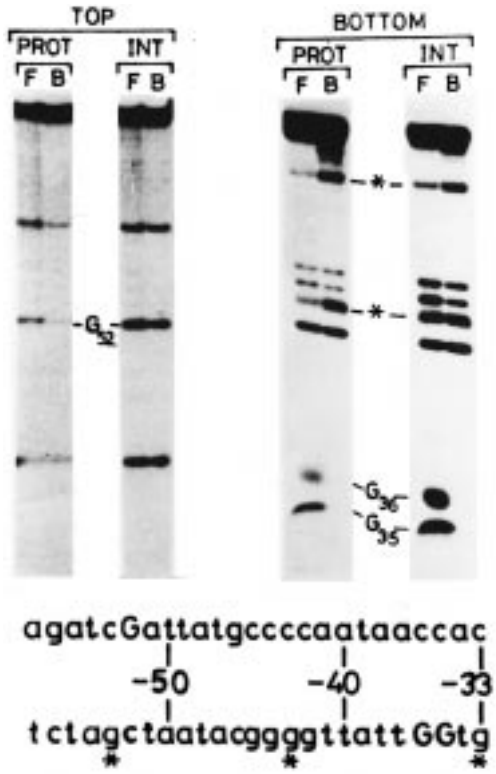

FIGURE 9: Purine contacts of C protein. The details of the experimental procedure are given in Materials and Methods. The DMS protection (PROT) and interference (INT) of guanines in the presence of C protein from top and bottom strands as indicated. F indicates free DNA; B indicates the DNA bound to the protein. The protected guanines are in uppercase, and asterisks indicate the hypermethylated guanines in the sequence depicted below the autorad. The guanine-33 hyperreactivity could not be detected with this substrate, but inferred using a larger DNA fragment (*12*).

changes in its CD behavior in the presence of the protein, suggesting that the C protein brings about the conformational changes only in its cognate site by specific interaction.

To substantiate the results with methyl green competition experiments which indicated the major groove interaction of the C protein, methylation protection interference analysis was carried out. Since the N7 position of guanine is located in the major groove of B-DNA (*23*), protection of this group from methylation by dimethyl sulfate (DMS) in the presence of DNA binding protein could indicate the nature of interaction. Methylation of DNA at the N7 position of guanine in saturating amounts of C protein was investigated by reacting the preformed protein−DNA complex with DMS. Residues −52 G in the top strand and −36 G and −35 G in the bottom strand were protected against methylation, indicating the proximity of these N7 groups to C protein in the complex (Figure 9). Since N7 guanine is a major groove determinant, the protection of these groups would mean a major groove interaction by the protein at these positions. Methylation interference experiments, wherein DNA is methylated with DMS before addition of the protein, support these observations. Interestingly, three guanine residues (−33 G, −43 G, and −53G) in the bottom strand are hyper-reactive, suggesting distortion of DNA. The hyper-reactive residues are located 10 nucleotides apart, reflecting axial distortion in DNA induced as a result of protein binding.

DISCUSSION

From the elaborate literature, it is apparent that transcriptional regulatory proteins have an array of structural motifs for DNA binding and use diverse mechanisms for transcription activation. In this paper, the contributions from the secondary structure of transcription activator C protein to its specific DNA binding have been addressed. The influence of various factors, viz., electrolytes and minor and major groove binders, on C protein−DNA complex formation have been studied to evaluate the DNA binding mechanism of the C protein.

The majority of the site-specific DNA binding proteins predominantly recognize the specific sequences through the major groove by hydrogen bonding with bases and electrostatic interactions with phosphate groups. However, there are examples such as TATA binding protein which bind sequence specifically to the minor groove of DNA (*24*). We have attempted to evaluate whether C protein binds to the minor or major groove of DNA. While methyl green competition experiments indicate the interaction of C protein at the major groove of DNA, DMS protection−interference analyses confirm this observation. On the other hand, the results obtained with distamycin competition experiments are intriguing and suggest two alternate possibilities for distamycin displacement from the minor groove; the conformational changes in DNA brought about upon C protein binding in the major groove could have inhibited the binding of distamycin to the opposite minor groove. An alternative explanation is that while the helix−turn−helix motif of C protein binds to DNA major groove the protein could make additional contacts with the minor groove, resulting in displacement of distamycin. An example for this mode of interaction is in the case of antennapedia homeodomain wherein additional contacts with the minor groove are made through a polypeptide loop immediately preceding the helix−turn−helix segment (*25*). Solution structure analysis of the C protein−DNA complex is needed to address this point.

Many enzymes require Mg^{2+} for their activity, and there are two general mechanisms for such an activation (*26*). One is through binding to a ligand, thereby making it a suitable substrate, e.g., ATP−Mg^{2+}, DNA−Mg^{2+}. The other mechanism is through allosteric activation in which binding to one more form of the enzyme in the reaction course or another site in the enzyme induces a change in its conformation. The latter seems to be the one applicable to the Mg^{2+}-mediated binding of C protein to its cognate DNA site. This is based on our CD results which show that Mg^{2+} brings about conformational changes in C protein from a more β-like structure to a more α-helical structure through allosteric activation (Figure 6). This β-helical structure, induced as a result of Mg^{2+} binding, seems to be a prerequisite for DNA recognition. Recently, Mg^{2+} has been shown to be the essential activator for protein tyrosine kinase (*27*). The activation mechanism involves the conformational changes in the enzyme induced by binding of Mg^{2+} either directly to the active site or to the sites other than the active site of the enzyme. In Mg^{2+}-dependent DNA−protein interactions, the binding of the protein to its cognate site may or may not be metal ion dependent. The role of Mg^{2+} can be best illustrated by taking the examples of two restriction enzymes, *Eco*RI and *Eco*RV. *Eco*RI can bind

DNA in absence of Mg^{2+}, but the divalent metal ion is required for the cleavage at the recognition sequence (*28*). On the other hand, Mg^{2+} is a prerequisite for the binding of *Eco*RV to its cognate site apart from the requirement for cleavage (*29, 30*). In the latter case, the presence of a second Mg^{2+} binding site other than the one in the catalytic center has been demonstrated. We have shown here that C protein binds to DNA with high affinity only in the presence of Mg^{2+}. Such high-affinity binding involves associated conformational changes in DNA, as demonstrated by CD studies (data not shown). This is supported by the presence of hypersites upon protein binding as revealed by the methylation protection−interference studies.

Zubay and Doty (*31*) noted that the α-helix would fit snugly into the B-DNA major groove. Others suggested that an antiparallel β-pleated sheet has the correct twist and helical structure for interaction with duplex (double-stranded) RNA (*32*) and DNA (*33*). It was shown by model building that an antiparallel β-ribbon (equivalent to a two-strand sheet) would fit into the B-DNA minor groove (*33*). The present CD observation confirms that the Mg^{2+} brings about conformational transitions from a more β-like structure to a more α helix structure which seems to be the essential requirement for major groove binding. Central to all the models for the interaction between certain repressor proteins and their operators is a pair of symmetry-related α-helixes, in a helix−turn−helix motif, that penetrate successive turns of the major groove. A search for putative DNA binding motifs in the 140 amino acid residues of C protein showed the presence of a helix−turn−helix (HTH) motif, spanning residues 100−119, in the carboxyl-terminal region of the protein (*7*). Our recent experiments with deletion mutants of C protein show that this region is important for DNA binding (*13*). A carboxyl terminal deletion spanning up to residue 113 abolished DNA binding implicating the role for the HTH motif in sequence recognition. Mg^{2+} could be stabilizing this helix−turn−helix motif by increasing α-helical content of C protein and thus providing specific geometrical properties and dimensions, complementary to the major groove of DNA helix. Such a molecular conformational switch could be the important step of the signal transduction in the transcriptional process.

ACKNOWLEDGMENT

We thank the anonymous referees for their constructive comments and help in improving the manuscript. We thank the Molecular Biology Unit, Banaras Hindu University for the CD facility and D. R. Radha for excellent technical assistance.

NOTE ADDED IN PROOF

Sun et al. (*34*) have also recently demonstrated C-mediated conformational changes in DNA.

REFERENCES

1. Beabealashivily, R. S., Ivanov, V. I., Minchenkova, L. E., and Savotchkina, L. P. (1972) *Biochim. Biophys. Acta 259*, 35−40.
2. Frederick, C. A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.-C., Greence, P., Boyer, H. W., and Rosenberg, J. M. (1984) *Nature (London) 309*, 327−331.
3. Vinson, C. R., Sigler, P. B., and McKnight, S. L. (1989) *Science 246*, 911−916.
4. Hattman, S., Ives, J., Margolin, W., and Howe, M. M. (1985) *Gene 39*, 71−76.
5. Heisig, P., and Kahmann, R. (1986) *Gene 43*, 59−67.
6. Hattman, S., Newman, L., Krishna Murthy, H. M., and Nagaraja, V. *Proc. Natl. Acad. Sci. U.S.A. 88*, 10027−10031.
7. Bolker, M., Wulczyn, F. G., and Kahmann, R. (1989) *J. Bacteriol. 171*, 2019−2027.
8. Margolin, W., and Howe, M. M. (1990) *J. Bacteriol. 172*, 1424−1429.
9. Balke, V., Nagaraja, V., Gindlesperger, T., and Hattman, S. (1992) *Nucleic Acids Res. 20*, 2777−2784.
10. Ramesh, V., De, A., and Nagaraja, V. (1994) *Protein Eng. 7*, 1053−1057.
11. Ramesh, V., De, A., and Nagaraja, V. (1994) *Protein Expression Purif. 5*, 379−384.
12. Ramesh, V., and Nagaraja, V. (1996) *J. Mol. Biol. 260*, 22−33.
13. De, A., Paul, B. D., Ramesh, V., and Nagaraja, V. (1997) *Protein Eng. 10*, 935−941.
14. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, (a) Chapter 1, pp 21−40, and (b) Chapter 11, pp 20−44, Cold Spring Harbor Laboratory Press, New York.
15. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol. 65*, 494−560.
16. Khadake, J. R., and Rao, M. R. S. (1995) *Biochemistry 34*, 15792−15801.
17. Bradford, M. (1976) *Anal. Biochem. 72*, 248−254.
18. Schmid, F. X. (1990) in *Protein Structure* (Creighton, T. E., Ed.) Chapter 11, pp 265−268, IRL Press, Oxford, U.K.
19. Adler, A. J., Greenfield, N. J., and Fasman, G. D. (1973) *Methods Enzymol. 27*, 697−735.
20. Chou, P. Y., and Fasman, G. D. (1978) *Adv. Enzymol. 47*, 45−147.
21. Zimmer, C., and Wahnert, U. (1986) *Prog. Biophys. Mol. Biol. 47*, 31−112.
22. Dasagupta, D., Parrack, P., and Sasisekharan, V. (1987) *Biochemistry 26*, 6381−6386.
23. Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804−808.
24. Lang, D., and Stamminger, T. (1994) *Nucleic Acids Res. 22*, 3331−3338.
25. Otting, G., Qian, Y. Q., Billeter, M., Muller, M., Affolter, M., Gehring, W. J., and Wuthrich, K. (1990) *EMBO J. 9*, 3085−3092.
26. Malmstrom, B. G., and Rosenberg, A. (1959) *Adv. Enzymol. Relat. Subj. Biochem. 21*, 131−167.
27. Sun, G., and Budde, R. J. A. (1997) *Biochemistry 36*, 2139−2146.
28. Vipond, I. B., Baldwin, G. S., and Halford, S. E. (1995) *Biochemistry 34*, 697−704.
29. Jeltsch, A., Maschke, H., Selent, U., Wenz, C., Kohler, E., Connolly, A., Throgood, H., and Pingoud, A. (1995) *Biochemistry 34*, 6239−6246.
30. Kostrewa, D., and Winkler, F. K. (1995) *Biochemistry 34*, 683−696.
31. Zubay, G., and Doty, P. (1959) *J. Mol. Biol. 1*, 1−20.
32. Carter, C. W., Jr., and Kraut, J. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*,283−287.
33. Church, G. M., Sussman, J. L., and Kim, S.−H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 1458−1462.
34. Sun, W., Hattman, S., and Kool, E. (1997) *J. Mol. Biol. 273*, 765−774.

BI972171V